CL #2

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/249,671

DATE: 06/16/94
TIME: 10:48:56

INPUT SET: S2395.raw

┌───┐
│ This Raw Listing contains the General   │
│ Information Section and those Sequences   │
│ containing ERRORS.   │
└───┘

wrapped text

```
 1                                  SEQUENCE LISTING
 2
 3    (1)     General Information:
 4
 5          (i)  APPLICANTS:       Start response on this line
 6                                  Hauptmann, R.
 7                                  Falkner, E.
 8                                  Bodo, G.
 9                                  VoÃ, T.
10                                  Maurer- Fogy, I.
11
12         (ii) TITLE OF INVENTION:  Process for Preparing and Purifying
13                                   alpha-Interferon
14
15        (iii) NUMBER OF SEQUENCES: 12
16
17         (iv) CORRESPONDENCE ADDRESS:
18              (A) ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
19              (B) STREET:  1100 New York Avenue, Suite 600
20              (C) CITY:  Washington
21              (D) STATE:  D.C.
22              (E) COUNTRY:  U.S.A.
23              (F) ZIP:  20005
24
25          (v) COMPUTER READABLE FORM:
26              (A) MEDIUM TYPE: Floppy disk
27              (B) COMPUTER: IBM PC compatible
28              (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29              (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
30
31         (vi) CURRENT APPLICATION DATA:
32              (A) APPLICATION NUMBER:
33              (B) FILING DATE:  herewith
34    Delete space  (C) CLASSIFICATION:
35                                  Information:
36       (viii) ATTORNEY/AGENT NUMBER:
37              (A) NAME:  Esmond, Robert W. RWE/EG
38              (B) REGISTRATION NUMBER:  32,893
39              (C) REFERENCE/DOCKET NUMBER: 0652.1350000
40
41         (ix) TELECOMMUNICATION INFORMATION:
42              (A) TELEPHONE:  (202) 371-2600
43              (B) TELEFAX:  (202) 371-2540
44
45
```
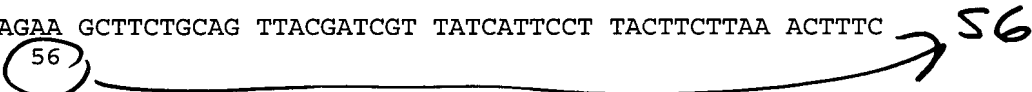
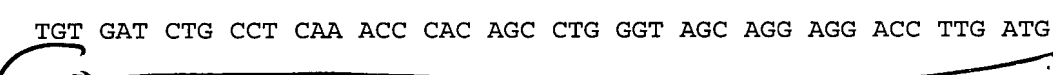*INPUT SET: S2395.raw*

46

ERRORED SEQUENCES FOLLOW:

```
61    (2) INFORMATION FOR SEQ ID NO: 2:
62
63         (i) SEQUENCE CHARACTERISTICS:
64             (A) LENGTH: 56 base pairs
65             (B) TYPE: nucleic acid
66             (C) STRANDEDNESS: both
67             (D) TOPOLOGY: unknown
68
69         (ii) MOLECULE TYPE: cDNA
70
71
72
73         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
74
75    GACTTCAGAA GCTTCTGCAG TTACGATCGT TATCATTCCT TACTTCTTAA ACTTTC
76         56
77
```

56

```
156   (2) INFORMATION FOR SEQ ID NO: 6:
157
158        (i) SEQUENCE CHARACTERISTICS:
159            (A) LENGTH: 495 base pairs
160            (B) TYPE: nucleic acid
161            (C) STRANDEDNESS: both
162            (D) TOPOLOGY: unknown
163
164        (ii) MOLECULE TYPE: cDNA
165
166
167        (ix) FEATURE:
168            (A) NAME/KEY: CDS
169            (B) LOCATION: 1..495
170
171
172        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
173
174   TGT GAT CTG CCT CAA ACC CAC AGC CTG GGT AGC AGG AGG ACC TTG ATG
175
176   48
177   Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met
178    1               5                   10                  15
179
180   CTC CTG GCA CAG ATG AGG AGA ATC TCT CTT TTC TCC TGC TTG AAG GAC
181
182   96
183   Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp
```

48

etc —

INPUT SET: S2395.raw

```
184                  20                          25                          30
185
186   AGA CGT GAC TTT GGA TTT CCC CAG GAG GAG TTT GGC AAC CAG TTC CAA
187
188   144
189   Arg Arg Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln
190            35                          40                          45
191
192   AAG GCT GAA ACC ATC CCT GTC CTC CAT GAG ATG ATC CAG CAG ATC TTC
193
194   192
195   Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile Phe
196         50                          55                          60
197
198   AAT CTC TTC AGC ACA AAG GAC TCA TCT GCT GCT TGG GAT GAG ACC CTC
199
200   240
201   Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr Leu
202    65                          70                          75            80
203
204   CTA GAC AAA TTC TAC ACT GAA CTC TAC CAG CAG CTG AAT GAC CTG GAA
205
206   288
207   Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu
208                  85                          90                          95
209
210   GCC TGT GTG ATA CAG GGG GTG GGG GTG ACA GAG ACT CCC CTG ATG AAG
211
212   336
213   Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met Lys
214                 100                         105                         110
215
216   GAG GAC TCC ATT CTG GCT GTG AGG AAA TAC TTC CAA AGA ATC ACT CTC
217
218   384
219   Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr Leu
220                 115                         120                         125
221
222   TAT CTG AAA GAG AAG AAA TAC AGC CCT TGT GCC TGG GAG GTT GTC AGA
223
224   432
225   Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg
226         130                         135                         140
227
228   GCA GAA ATC ATG AGA TCT TTT TCT TTG TCA ACA AAC TTG CAA GAA AGT
229
230   480
231   Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu Ser
232    145                         150                         155            160
233
234   TTA AGA AGT AAG GAA
235   495
236   Leu Arg Ser Lys Glu
```

```
        237                    165
        238
        239
```

```
        284    (2)  INFORMATION FOR SEQ ID NO: 8:
        285
        286           (i)  SEQUENCE CHARACTERISTICS:
-->     287               (A) LENGTH: 869 base pairs
        288               (B)  TYPE: nucleic acid
        289               (C)  STRANDEDNESS: both
        290               (D)  TOPOLOGY: unknown
        291
        292          (ii)  MOLECULE TYPE: cDNA
        293
        294
        295
        296          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 8:
        297
        298    GAATTCGAGA  TTATCGTCAC  TGCAATGCTT  CGCAATATGG  CGCAAAATGA
-->     299    CCAACAGCGG       60
        300
        301    TTGATTGATC  AGGTAGAGGG  GGCGCTGTAC  GAGGTAAAGC  CCGATGCCAG
        302    CATTCCTGAC
        303    120
        304
        305    GACGATACGG  AGCTGCTGCG  CGATTACGTA  AAGAAGTTAT  TGAAGCATCC
        306    TCGTCAGTAA
        307    180
        308
        309    AAAGTTAATC  TTTTCAACAG  CTGTCATAAA  GTTGTCACGG  CCGAGACTTA
        310    TAGTCGCTTT
        311    240
        312
        313    GTTTTTATTT  TTTAATGTAT  TTGCTCGAGA  GGTTGAGGTG  ATTTTATGAA
        314    AAAGAATATC
        315    300
        316
        317    GCATTTCTTC  TTGCATCTAT  GTTCGTTTTT  TCTATTGCTA  CAAATGCCTA
        318    TGCATGTGAT
        319    360
        320
        321    CTGCCTCAAA  CCCACAGCCT  GGGTAGCAGG  AGGACCTTGA  TGCTCCTGGC
        322    ACAGATGAGG
        323    420
        324
        325    AGAATCTCTC  TTTTCTCCTG  CTTGAAGGAC  AGACGTGACT  TTGGATTTCC
        326    CCAGGAGGAG
        327    480
        328
        329    TTTGGCAACC  AGTTCCAAAA  GGCTGAAACC  ATCCCTGTCC  TCCATGAGAT
        330    GATCCAGCAG
        331    540
        332
```

INPUT SET: S2395.raw

```
333   ATCTTCAATC TCTTCAGCAC AAAGGACTCA TCTGCTGCTT GGGATGAGAC
334   CCTCCTAGAC
335   600
336
337   AAATTCTACA CTGAACTCTA CCAGCAGCTG AATGACCTGG AAGCCTGTGT
338   GATACAGGGG
339   660
340
341   GTGGGGGTGA CAGAGACTCC CCTGATGAAG GAGGACTCCA TTCTGGCTGT
342   GAGGAAATAC
343   720
344
345   TTCCAAAGAA TCACTCTCTA TCTGAAAGAG AAGAAATACA GCCCTTGTGC
346   CTGGGAGGTT
347   780
348
349   GTCAGAGCAG AAATCATGAG ATCTTTTTCT TTGTCAACAA ACTTGCAAGA
350   AAGTTTAAGA
351   840
352
353   AGTAAGGAAT GATAACGATC GTAACTGCA
354   869
355
```

```
356   (2)  INFORMATION FOR SEQ ID NO: 9:
357
358        (i)  SEQUENCE CHARACTERISTICS:
359             (A)  LENGTH: 1177 base pairs
360             (B)  TYPE: nucleic acid
361             (C)  STRANDEDNESS: both
362             (D)  TOPOLOGY: unknown
363
364       (ii)  MOLECULE TYPE: cDNA
365
366
367       (ix)  FEATURE:
368             (A)  NAME/KEY: CDS
369             (B)  LOCATION: 286..873
370             (D)  OTHER INFORMATION: /function= "Cytokine"
371                       /product= "Interferon-omega1"
372
373       (ix)  FEATURE:
374             (A)  NAME/KEY: mat_peptide
375             (B)  LOCATION: 355..873
376             (D)  OTHER INFORMATION: /function= "Cytokine"
377                       /product= "Interferon-omega"
378
379       (ix)  FEATURE:
380             (A)  NAME/KEY: sig_peptide
381             (B)  LOCATION: 286..354
382             (D)  OTHER INFORMATION: /product= "ST II Leader"
383
384
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/249,671*

INPUT SET: S2395.raw

```
385        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
386
387   GAATTGGAGA TTATCGTCAC TGCAATGCTT CGCAATATGG CGCAAAATGA
388   CCAACAGCGG
389   60
390
391   TTGATTGATC AGGTAGAGGG GGCGCTGTAC GAGGTAAAGC CCGATGCCAG
392   CATTCCTGAC
393   120
394
395   GACGATACGG AGCTGCTGCG CGATTACGTA AAGAAGTTAT TGAAGCATCC
396   TCGTCAGTAA
397   180
398
399   AAAGTTAATC TTTTCAACAG CTGTCATAAA GTTGTCACGG CCGAGACTTA
400   TAGTCGCTTT
401   240
402
403   GTTTTTATTT TTTAATGTAT TTGCTCGAGA GGTTGAGGTG ATTTT ATG AAA AAG
404
405   294
406                                                 Met Lys Lys
407                                                 -23
408
409   AAT ATC GCA TTT CTT CTT GCA TCT ATG TTC GTT TTT TCT ATT GCT ACA
410   342
411   Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser Ile Ala Thr
412   -20               -15             -10                     -5
413
414   AAT GCC TAT GCA TGT GAT CTG CCT CAG AAC CAT GGC CTA CTT AGC AGG
415
416   390
417   Asn Ala Tyr Ala Cys Asp Leu Pro Gln Asn His Gly Leu Leu Ser Arg
418                     1           5                       10
419
420   AAC ACC TTG GTG CTT CTG CAC CAA ATG AGG AGA ATC TCC CCT TTC TTG
421
422   438
423   Asn Thr Leu Val Leu Leu His Gln Met Arg Arg Ile Ser Pro Phe Leu
424            15                20                25
425
426   TGT CTC AAG GAC AGA AGA GAC TTC AGG TTC CCC CAG GAG ATG GTA AAA
427
428   486
429   Cys Leu Lys Asp Arg Arg Asp Phe Arg Phe Pro Gln Glu Met Val Lys
430         30                35                40
431
432   GGG AGC CAG TTG CAG AAG GCC CAT GTC ATG TCT GTC CTC CAT GAG ATG
433
434   534
435   Gly Ser Gln Leu Gln Lys Ala His Val Met Ser Val Leu His Glu Met
436      45                50                55                60
437
```

INPUT SET: S2395.raw

```
438   CTG CAG CAG ATC TTC AGC CTC TTC CAC ACA GAG CGC TCC TCT GCT GCC
439
440   582
441   Leu Gln Gln Ile Phe Ser Leu Phe His Thr Glu Arg Ser Ser Ala Ala
442                   65              70              75
443
444   TGG AAC ATG ACC CTC CTA GAC CAA CTC CAC ACT GGA CTT CAT CAG CAA
445
446   630
447   Trp Asn Met Thr Leu Leu Asp Gln Leu His Thr Gly Leu His Gln Gln
448                   80              85              90
449
450   CTG CAA CAC CTG GAG ACC TGC TTG CTG CAG GTA GTG GGA GAA GGA GAA
451
452   678
453   Leu Gln His Leu Glu Thr Cys Leu Leu Gln Val Val Gly Glu Gly Glu
454           95              100             105
455
456   TCT GCT GGG GCA ATT AGC AGC CCT GCA CTG ACC TTG AGG AGG TAC TTC
457
458   726
459   Ser Ala Gly Ala Ile Ser Ser Pro Ala Leu Thr Leu Arg Arg Tyr Phe
460       110             115             120
461
462   CAG GGA ATC CGT GTC TAC CTG AAA GAG AAG AAA TAC AGC GAC TGT GCC
463
464   774
465   Gln Gly Ile Arg Val Tyr Leu Lys Glu Lys Lys Tyr Ser Asp Cys Ala
466   125             130             135             140
467
468   TGG GAA GTT GTC AGA ATG GAA ATC ATG AAA TCC TTG TTC TTA TCA ACA
469
470   822
471   Trp Glu Val Val Arg Met Glu Ile Met Lys Ser Leu Phe Leu Ser Thr
472               145             150             155
473
474   AAC ATG CAA GAA AGA CTG AGA AGT AAA GAT AGA GAC CTG GGC TCA TCT
475
476   870
477   Asn Met Gln Glu Arg Leu Arg Ser Lys Asp Arg Asp Leu Gly Ser Ser
478               160             165             170
479
480   TGAAATGATT CTCATTGATT AATTTGCCAT ATAACACTTG CACATGTGAC
481   TCTGGTCAAT
482   930
483
484   TCAAAGACT CTTATTTCGG CTTTAATCAC AGAATTGACT GAATTAGTTC
485   TGCAAATACT
486   990
487
488   TTGTCGGTAT ATTAAGCCAG TATATGTTAA AAAGACTTAG GTTCAGGGGC
489   ATCAGTCCCT
490   1050
```

INPUT SET: S2395.raw

```
491
492    AAGATGTTAT TTATTTTTAC TCATTTATTT ATTCTTACAT TTTATCATAT
493    TTATACTATT
494    1110
495
496    TATATTCTTA TATAACAAAT GTTTGCCTTT ACATTGTATT AAGATAACAA
497    AACATGTTCA
498    1170
499
500    GGATCCA
501    1177
502
```

```
553    (2)  INFORMATION FOR SEQ ID NO: 11:
554
555         (i)  SEQUENCE CHARACTERISTICS:
556              (A)  LENGTH: 879 base pairs
557              (B)  TYPE: nucleic acid
558              (C)  STRANDEDNESS: both
559              (D)  TOPOLOGY: unknown
560
561        (ii)  MOLECULE TYPE: cDNA
562
563
564
565        (ix)  FEATURE:
566              (A)  NAME/KEY: CDS
567              (B)  LOCATION: 286..852
568
569        (ix)  FEATURE:
570              (A)  NAME/KEY: mat_peptide
571              (B)  LOCATION: 355..852
572              (D)  OTHER INFORMATION: /function= "Cytokine"
573                   /product= "Interferon-alpha-2c"
574
575        (ix)  FEATURE:
576              (A)  NAME/KEY: sig_peptide
577              (B)  LOCATION: 286..354
578              (D)  OTHER INFORMATION: /product= "ST II Leader"
579
580
581        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 11:
582
583    GAATTCGAGA TTATCGTCAC TGCAATGCTT CGCAATATGG CGCAAAATGA
584    CCAACAGCGG
585    60
586
587    TTGATTGATC AGGTAGAGGG GGCGCTGTAC GAGGTAAAGC CCGATGCCAG
588    CATTCCTGAC
589    120
590
591    GACGATACGG AGCTGCTGCG CGATTACGTA AAGAAGTTAT TGAAGCATCC
592    TCGTCAGTAA
```

PAGE: 9

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/249,671

DATE: 06/16/94
TIME: 10:50:05

INPUT SET: S2395.raw

```
593    180
594
595    AAAGTTAATC TTTTCAACAG CTGTCATAAA GTTGTCACGG CCGAGACTTA
596    TAGTCGCTTT
597    240
598
599    GTTTTTATTT TTTAATGTAT TTGCTCGAGA GGTTGAGGTG ATTTT ATG AAA AAG
600
601    294
602                                                    Met Lys Lys
603                                                    -23
604
605    AAT ATC GCA TTT CTT CTT GCA TCT ATG TTC GTT TTT TCT ATT GCT ACA
606    342
607    Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe Ser Ile Ala Thr
608    -20            -15              -10                       -5
609
610    AAT GCC TAT GCA TGT GAT CTG CCT CAA ACC CAC AGC CTG GGT AGC AGG
611
612    390
613    Asn Ala Tyr Ala Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg
614                    1            5                     10
615
616    AGG ACC TTG ATG CTC CTG GCA CAG ATG AGG AGA ATC TCT CTT TTC TCC
617
618    438
619    Arg Thr Leu Met Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser
620            15                  20                  25
621
622    TGC TTG AAG GAC AGA CGT GAC TTT GGA TTT CCC CAG GAG GAG TTT GGC
623
624    486
625    Cys Leu Lys Asp Arg Arg Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly
626         30                  35                  40
627
628    AAC CAG TTC CAA AAG GCT GAA ACC ATC CCT GTC CTC CAT GAG ATG ATC
629
630    534
631    Asn Gln Phe Gln Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile
632     45                  50                  55                  60
633
634    CAG CAG ATC TTC AAT CTC TTC AGC ACA AAG GAC TCA TCT GCT GCT TGG
635
636    582
637    Gln Gln Ile Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp
638                    65                  70                  75
639
640    GAT GAG ACC CTC CTA GAC AAA TTC TAC ACT GAA CTC TAC CAG CAG CTG
641
642    630
643    Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu
644                    80                  85                  90
645
```

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/08/249,671*

INPUT SET: S2395.raw

```
646    AAT GAC CTG GAA GCC TGT GTG ATA CAG GGG GTG GGG GTG ACA GAG ACT
647
648    678
649    Asn Asp Leu Glu Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr
650            95                100               105
651
652    CCC CTG ATG AAG GAG GAC TCC ATT CTG GCT GTG AGG AAA TAC TTC CAA
653
654    726
655    Pro Leu Met Lys Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln
656        110                115               120
657
658    AGA ATC ACT CTC TAT CTG AAA GAG AAG AAA TAC AGC CCT TGT GCC TGG
659
660    774
661    Arg Ile Thr Leu Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp
662    125               130               135               140
663
664    GAG GTT GTC AGA GCA GAA ATC ATG AGA TCT TTT TCT TTG TCA ACA AAC
665
666    822
667    Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn
668                145               150               155
669
670    TTG CAA GAA AGT TTA AGA AGT AAG GAA TGATAACGAT CGTAACTGCA
671
672    869
673    Leu Gln Glu Ser Leu Arg Ser Lys Glu
674                160               165
675
676    GAAGCTTAAT
677    879
678
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/249,671*

INPUT SET: S2395.raw

Line	Error	Original Text
5	Mandatory Value Not Present	(i) APPLICANTS:
36	Unknown or Misplaced Identifier	(viii) ATTORNEY/ AGENT NUMBER:
37	Unknown or Misplaced Identifier	(A) NAME: Esmond, Robert W. RWE/EG
38	Unknown or Misplaced Identifier	(B) REGISTRATION NUMBER: 32,893
39	Unknown or Misplaced Identifier	(C) REFERENCE/DOCKET NUMBER: 0652.1350000
64	Entered (56) and Calc. Seq. Length (0) differ	(A) LENGTH: 56 base pairs
159	Entered (495) and Calc. Seq. Length (0) differ	(A) LENGTH: 495 base pairs
287	Entered (869) and Calc. Seq. Length (10) differ	(A) LENGTH: 869 base pairs
299	# of Sequences for line conflicts w/ running total	CCAACAGCGG 60
359	Entered (1177) and Calc. Seq. Length (0) differ	(A) LENGTH: 1177 base pairs
556	Entered (879) and Calc. Seq. Length (0) differ	(A) LENGTH: 879 base pairs